

# 在线教育学习者学习有效性评价及差异性研究

索 琪, 左 佩

(青岛科技大学 经济与管理学院, 山东 青岛 266061)

**摘要:**在线教育具有开放性和共享性特征,深入挖掘在线学习者学习行为,能够有助于学生及时了解学习效果。选取在线教育平台思政类、外语类和管理类课程为对象,研究在线学习者的学习有效性及群体差异性特征。研究表明,不同类型课程的学习者行为和成绩分布具有相似性,大部分学习者并未实现高效的学习效率,学习行为多且成绩较好的“勤奋收获者”与“学习体验者”学习效率之间存在明显的差异性。基于此,提出在线教育的过程中,应当加强阶段性监督,对缺乏学习积极性的学生进行及时干预,完善线上教育的筛选制度等措施,以期提高学习者的效率水平和发挥在线课程的最大利用程度。

**关键词:**在线学习行为;聚类分析;数据包络分析;效率提升策略

**中图分类号:**G434

**文献标识码:**A

**文章编号:**2097-0625(2024)01-0038-09

## 一、引言

长期以来,教育资源分布不均衡、时空约束等问题严重制约了传统教育模式下教育质量提升以及区域人才的培育发展。随着互联网的普及和发展,“在线教育”这一新兴教育模式扩展了传统线下授课方式,是线下教学模式的重要补充,在线教育成为“互联网+教育”重要载体的同时,要求更高层次的监管规范以形成教育新秩序<sup>[1]</sup>。2020 年新冠肺炎疫情暴发以来,大量的线下课程转为线上形式,推动着在线教育市场的迅速发展,在线学习者的“近期价值”“自主感”和“工具性”也成为重要的需求驱动力<sup>[2]</sup>。据《2022 年度中国数字教育市场数据报告》<sup>[3]</sup>显示,2022 年中国数字教育市场规模 3 620 亿元,在线教育行业的用户规模达 3.14 亿人,高等教育借助在线平台发展的模式已成趋势。

在线教育平台在学习者使用过程中,记录了大量的学习数据。通过对数据分析发现,开放性网络环境下,学习者存在着学习资源利用率低<sup>[4]</sup>、师生之间缺乏互动<sup>[5]</sup>以及配套支持政策不到位<sup>[6]</sup>等一系列问题。在线学习形式与传统学习形式存在着本质性差异,如何有效地挖掘、分析并利用在线学习产生的大量数据,从而跟踪学习者的学习状态,适时调整教学方式

和侧重点,学习者如何根据在线行为数据对自己的学习水平进行准确的评估,逐渐成为研究者关注的焦点问题。通过充分挖掘学习行为大数据并进行统计分析,掌握学习者在线学习行为特征和学习偏好,识别影响学习行为的关键因素,获得学习者群体学习行为规律,并提出学习支持服务策略,能够为提升教学质量提供数据支撑和决策参考<sup>[7]</sup>。

以往传统的在线学习行为研究主要基于统计分析的相关方法,王佳利和李斌峰证明了混合教学模式中,学习者的成绩明显优于传统教学模式,认为良好的混合式教学模式能够提升教学质量和学生满意度<sup>[8]</sup>。随着研究的深入,学者将机器学习算法等数据挖掘方法引入研究中,胡祖辉和施佳采用数据挖掘方法将学生上网行为与成绩进行相关性分析,发现上网行为对于成绩具有负面影响,提出制定相应的网络管理措施从而限制学生过度上网<sup>[9]</sup>。贺超凯和吴蒙采用逻辑回归算法对课程成绩进行预测,证明学生教育背景和性别对于学习效果影响不大,通过学习行为能够准确预测学习效果<sup>[10]</sup>。近年来,研究对象逐渐由以学生个体为中心转向关注学习者的群体特征,Wang 和 Zhang 将 6 个学习行为指标作为解释变量,将学习者划分为注重教材学习而缺乏互动的一类学

**收稿日期:**2023-11-20

**基金项目:**中国高等教育学会高等教育科学研究规划课题重大项目“大数据背景下在线学习行为有效性评价”(项目编号:22SZH0101)

**作者简介:**索琪(1980—),女,吉林扶余人,教授,博士。研究方向:复杂系统建模、大数据挖掘。

习者、在线时间和互动能力都较好的二类学习者以及所有学习行为都缺乏主动性的三类学习者<sup>[11]</sup>。赵磊等发现学习行为和学习成绩之间存在显著相关性,对学习者的 11 项关键行为数据进行聚类分析,将所有学习者分为优秀学习者、良好学习者、普通学习者和风险学习者,针对四种学习者的不同行为特点提出了优化课程设计、创新教学模式和注重发展性评价的教学改进策略<sup>[12]</sup>。王改花和傅钢善基于数据挖掘和聚类分析发现,学习沉浸度较高的学习者的学习效果通常较好<sup>[13]</sup>。此外,当前研究从单纯学习行为数据分析转向结合学习效率的分析,Fuentes 等<sup>[14]</sup>建立三阶段 DEA 模型测算课程的学习效率,发现课程满意度、材料多样性和教师满意度是影响教学的重要因素。陈长胜等将聚类分析与 DEA 相结合,识别了慕课学习者的学习行为差异,并据此提出提高学习者的学情监控能力、以奖励激发学习动力等效率提升策略<sup>[15]</sup>。

综合国内外相关文献分析,目前对于学生学习行为的分析已经取得初步成果,对于学生学习效率的分析多采用 DEA 方法,研究方法已经比较成熟,但针对多课程的在线学习分析及群组差异性分析的相关研究较少。因此,本文为深入探究在线学习者的学习有效性、差异性、群组特征,在挖掘在线教育平台不同类型课程学习者学习行为数据基础上,采用聚类方法分析思政、外语和管理课程学习者学习行为的群组差异性,采用 DEA 方法对学生的学习效率进行分析,根据学习者的效率群组差异性提出改善学习效率的策略和建议,以期为在线教育平台的学习评价提供借鉴。

## 二、研究设计

### (一) 研究方法

#### 1. 聚类分析

聚类,是根据数据内部的结构特征将数据划分成不同簇的一种研究方法。通常通过定义距离和相似系数,将全部数据划分成若干组,使得组内数据具有相似性特征<sup>[16]</sup>,而组间数据则具有较大差异性。通过对数据进行划分,找寻数据之间隐含的规律,可以有效避免按照经验进行定性分类产生的主观随意性<sup>[17]</sup>。本文采用聚类分析方法参照学习行为数据特征对学习者的学习行为进行分组,作为后续学习效率对比分析和

效率提升研究的起点。

### 2. 数据包络分析

数据包络分析(Data Envelopment Analysis,简称 DEA)是一种基于决策对象之间比较评价的效率分析方法。DEA 分析在计算多个投入和产出时具有特殊优势,能够快速计算不同决策对象的多维度指标投入产出效率<sup>[18]</sup>。

基础 CCR 模型一般假定规模效率不变,其得出的技术效率包括了规模效率的成分,因此被称为“综合技术效率”。若综合技术效率为 1 则称决策单元有效,若综合技术效率小于 1 则称决策单元无效<sup>[19]</sup>。BCC 模型基于规模效率可变的假设,所得出的效率被称为“纯技术效率”“综合技术效率”与“纯技术效率”的商值即为“规模效率”。超效率模型中被评价的决策单元的效率根据其他有效的决策单元构成的前沿得出,从而使得有效的决策单元的效率大于 1,是对 BCC 模型的扩展与完善<sup>[20]</sup>。本文采用以上三种模型分别对学习者的学习综合效率、学习技术效率、学习规模效率和学习超效率进行测算。

假设有  $n$  个决策单元,各有  $m$  种投入  $x_i (i=1, 2, 3, \dots, m)$ , 对应  $r$  种产出  $y_j (j=1, 2, 3, \dots, r)$ 。 $s^-$ 、 $s^+$  分别表示投入和产出的松弛变量, $\lambda$  表示权重向量, $\rho$  表示学习超效率值, $\rho \geq 1$  表示决策单元的相对有效。超效率 SBM 模型如式(1)所示。

$$\rho = \min \frac{1 + \frac{1}{m} \sum_{i=1}^m \frac{s_i^-}{x_{ik}}}{1 - \frac{1}{r} \sum_{j=1}^r \frac{s_j^+}{y_{rk}}} \quad \text{s. t.} \begin{cases} \sum_{j=1, j \neq k}^n x_{ij} \lambda_j - s_i^- \leq x_{ik} \\ \sum_{j=1, j \neq k}^n y_{ij} \lambda_j - s_i^+ \geq x_{rk} \\ \lambda, s^-, s^+ \geq 0 \\ i = 1, 2, \dots, m; r = 1, 2, \dots, q; \\ j = 1, 2, \dots, n (j \neq k) \end{cases} \quad (1)$$

### (二) 研究对象及数据来源

结合数据规模和数据类型的多样性,本文选取某开放大学在线学习平台为实证分析对象,分别选取思政课程、外语课程和专业课程这三类课程进行分析。在对数据初步筛选时,基于以下原则选择代表性课程:一方面,要求课程的学习活动较为完整,以保证指标数据的完整性,从而保障模型计算的基本规范要

求;另一方面,要求课程参与学习的学习者数量多、代表性强,从而使得分析结论具有普适性特征。鉴于此,本文选取思想道德修养与法律基础、人文英语 3 以及管理学基础三门课程为分析对象,学期选取 2021 年度春季学期。通过 DEA 初步计算发现,少量学习者的学习行为数据不符合 CCR 模型和 BCC 模型的计算条件,即所有的产出指标都是 0,缺乏研究意义,因此将该部分学习数据剔除,最终三门课程的 DEA 决策单元数量依次为 1 618、835、1 081 人。

(三)指标选取

研究发现,在线学习行为指标能够刻画学习者的学习态度和学习参与程度<sup>[21]</sup>。龚艺等以出勤情况和资源浏览的行为数刻画学习者的参与程度,以完成作业的个数和频次刻画学习者的学习专注情况,以发帖数刻画学习过程中的交互情况<sup>[22]</sup>。沈欣忆等综合学

生论坛发帖行为、视频观看情况、学习时间偏好、作业完成情况以及学习页面的访问次数等指标构建学习行为度量模型<sup>[23]</sup>。本文根据平台记录的学习数据,选择刻画学生学习行为的投入指标包括:课程行为天数、行为总数、浏览数以及浏览资源数。李梅等在对课程开放度与成绩相关性的研究中,采用学习活动与最终成绩的加权之和刻画课程成绩<sup>[24]</sup>。胡祖辉和施佳引入了平均绩点指标,以满足数据挖掘的需要,减少成绩的数据复杂程度<sup>[9][28]</sup>。为进行 DEA 分析量化产出水平,本文选取试卷成绩、形考成绩(平时成绩)、最终成绩这三部分成绩作为产出指标,这三部分成绩指标综合了学习者平时的努力程度、学习情况和最终考试的成绩结果,能够较为客观地刻画出学习者的学习效果,最终构建的指标体系如表 1 所示。

表 1 DEA 指标体系

指标	指标名称	指标代号	指标计算
投入指标	课程行为天数	$X_1$	学习者参与本门课程进行学习行为的天数
	行为总数	$X_2$	与课程有关的学习行为总次数,包括课程点击率、完成作业、论坛发帖等
	浏览数	$X_3$	学习者在教学平台的浏览次数
	浏览资源数	$X_4$	学习者在教学平台浏览学习资源的次数
产出指标	试卷成绩	$Y_1$	最后考试的卷面成绩
	形考成绩	$Y_2$	学习过程中的平时成绩
	最终成绩	$Y_3$	由试卷成绩和形考成绩各占比 50% 计算得出

为挖掘学生行为模式和效率的组间差异,因此需要在综合 DEA 全部学习行为和成绩指标体系的基础上进行聚类分析,选取 DEA 分析选取的全部学习行为指标和学习成绩指标进行七维聚类,发现其分布上的相似性,通过对聚类组别分组效果的对比最后将每门课程的所有学习者分为两类。

三、研究结果及分析

(一)描述性分析

1. 学生学习行为分析

各门课程的描述性统计结果如表 2 所示。就三门课程整体而言,学习者各个指标之间均存在较大差异,所有学科均存在几乎没有学习行为的学习者,这使得每个指标组间的极差接近最大值,反映出学习

者对课程的重视程度不同。就不同类型课程而言,三类课程的平均行为天数差异较小,平均行为总数和浏览

表 2 学生学习行为统计分析

课程	指标	平均次数/次	极差	标准差
思政课程	课程行为天数	9.1	79	6.3
	行为总数	474.0	1 425	180.1
	浏览数	184.2	1 209	133.3
	浏览资源数	26.0	328	29.6
外语课程	课程行为天数	8.7	41	5.6
	行为总数	816.2	2 547	463.3
	浏览数	304.4	1 515	265.4
	浏览资源数	31.3	434	36.7
管理课程	课程行为天数	7.2	52	6.2
	行为总数	677.7	3 687	680.2
	浏览数	241.8	2 863	320.0
	浏览资源数	23.0	1 447	71.4



侧重点有所不同。思政课程的学习者相对行为总数较少,主要源于这门课程对于学习者平时进行网页学习互动的要求较少。平均行为总数和浏览次数较多的课程其极差和标准差也比较大,在学习过程中学习者的差别较为明显,同一门课程中学习者的学习方式也存在较大差异。三类课程中各项指标差别最为明显的为管理课程,表明该课程的学习者学习程度差别最大,专业课程与公共课程相比往往存在一定的难度,因此导致部分学习者无法跟上整体进度。

2. 学生成绩分析

三门课程的学生成绩统计信息如表 3 所示,大多数学习者顺利完成了全部学习过程,掌握了课程相关知识并顺利通过了该课程。

表 3 学生成绩统计分析

课程	课程通过比例	平均最终成绩	成绩优于平均值人数	成绩中位数
思政课程	90.3%	84.7	1 377	94
外语课程	89.3%	77.3	611	85
管理课程	79.8%	76.4	877	89

整体而言,课程通过率比较高,且平均成绩均在 75 分以上、中位数在 85 分以上,且在通过考试的学习者中 80% 以上成绩高于平均成绩。然而,三门课

程的频数分布存在差异。思政课程的频数最高分布区间是 90~100 分,而外语和管理课程的频数最高分布区间为 80~95 分。其中思政课程最终分值在 95~100 分的学习者占比高达 45%,该门课程的学习者更易取得好成绩。三门课程均存在低分者(<60 分)较多的情况,分别占比 9.70%、10.86%、16.46%,说明一定的学习者对课程不够重视,甚至可能存在弃考现象。整体频数分布在 65~95 分之间存在稳定的上升趋势,这符合在学习认真者中高分多、低分少的实际分布规律。

(二) 聚类分析

基于 Python 进行 K-means 聚类,结果如表 4 所示。聚类分析的轮廓系数是评估的重要指标,轮廓系数越大表示聚类的效果越好,通过比对不同分类数的轮廓系数证明某一类聚类效果最好。因此,将每门课程的学习者均分为两类,三门课程的轮廓系数均在 0.5 左右,表明聚类性能较好。不同类别的相关学习行为指标之间存在显著性差异,这表明学习者的学习行为具有明显差异性。思政课程成绩和学习行为数明显优于外语课程的特征,结合陈长胜等的研究<sup>[14]</sup>,将两类学习者分别定义为“勤奋收获者”与“学习体验者”。

表 4 学习行为的聚类分析结果及组间差异

课程	指标	勤奋收获者		学习体验者	
		均值	方差	均值	方差
思政课程轮廓系数=0.56	课程行为天数	9.47	40.96	5.18	17.75
	行为总数	493.95	27 498.25	288.01	40 439.35
	浏览数	191.97	17 807.25	111.90	11 753.14
	浏览资源数	26.87	899.98	17.68	558.12
	试卷成绩	89.10	96.83	24.27	497.35
	形考成绩	97.88	32.73	58.66	1 916.73
	最终成绩	93.69	30.88	0.84	37.30
外语课程轮廓系数=0.50	课程行为天数	9.08	32.27	5.25	15.09
	行为总数	856.73	202 246.11	471.14	189 469.04
	浏览数	325.22	71 283.20	127.41	28 630.82
	浏览资源数	32.70	1 410.75	19.60	665.12
	试卷成绩	72.21	158.24	14.61	769.60
	形考成绩	95.75	38.06	65.26	1 295.48
	最终成绩	84.21	51.23	18.02	527.91

续表 4 学习行为的聚类分析结果及组间差异

课程	指标	勤奋收获者		学习体验者	
		均值	方差	均值	方差
管理课程轮廓系数=0.48	课程行为天数	7.72	39.57	4.61	21.98
	行为总数	709.58	473 820.43	514.85	375 733.93
	浏览数	248.86	103 546.30	205.67	95 540.92
	浏览资源数	23.71	5 620.51	19.49	2 422.01
	试卷成绩	82.07	94.73	23.12	1 325.43
	形考成绩	94.39	46.84	51.15	1 948.81
	最终成绩	88.48	36.37	14.76	486.22

在对三门课程的两类学习者各项指标数据进行单因素方差检验时发现,所有的  $p$  值均小于 0.01,表明两类学习者的所有指标在统计学上均具有显著差异。这直观表明两类学习者从事各项学习活动的指标值差异明显,其中在最终成绩的指标上差异最为明显,勤奋收获者大部分通过了课程,而未通过课程的学习者则主要分布在学习体验者中,只有极少数的学习者没有满足该规律。聚类分析数据降维后的二维

表示图如图 1 所示,能够直观地显示出学习者之间呈现出分类聚集的特征,其中数量较多的集群为勤奋收获者,数量较少的集群为学习体验者。

(三)DEA 效率分析

对于每门课程,每一名学习者作为一个决策单元,将该课程所有学习者进行汇总用于 DEA 分析。三门课程的学习者效率统计情况如表 5 所示。

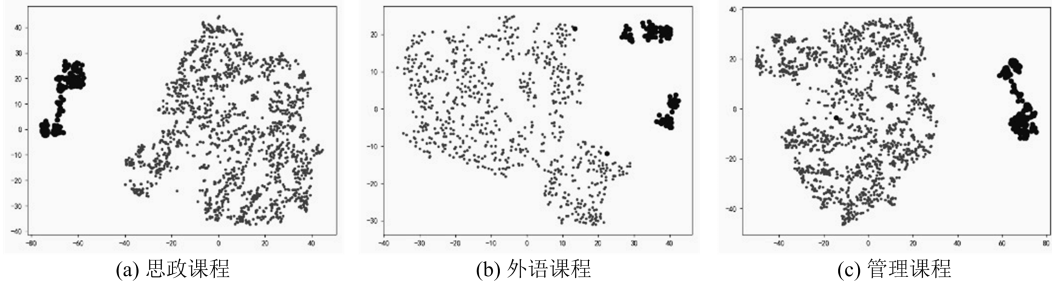


图 1 聚类分析二维表示图

表 5 课程学习效率统计情况对比

课程类别	效率指标	有效数量(人)	有效占比(%)	整体 $\bar{x}$	最小值	勤奋收获者 $\bar{x}$	学习体验者 $\bar{x}$
思政课程 N=1 618	综合效率	35	2.16	0.19	0.003	0.18	0.27
	纯技术效率	924	57.11	0.97	0.01	0.98	0.78
	规模效率	61	3.77	0.20	0.01	0.19	0.37
	超效率	24	1.48	0.88	0.27	0.93	0.28
外语课程 N=835	综合效率	14	1.68	0.20	0.03	0.18	0.25
	纯技术效率	191	22.87	0.95	0.18	0.95	0.79
	规模效率	21	2.52	0.22	0.03	0.20	0.39
	超效率	40	4.79	0.82	0.12	0.88	0.29
管理课程 N=1 081	综合效率	16	1.48	0.26	0.01	0.25	0.33
	纯技术效率	43	3.98	0.94	0.10	0.95	0.84
	规模效率	84	7.77	0.28	0.02	0.26	0.39
	超效率	26	2.41	0.78	0.04	0.89	0.13

1. 思政课程 DEA 分析结果

(1)学习效率分析

综合效率为 1 的学习者占比 2.16%,综合效率小于 0.5 的学习者占比达 93.45%,说明仅有极少数学习者实现了综合技术效率有效。在所有学习者中,规模报酬递减的人数占比为 86.40%,这类学习者按比例增加投入只能获得相对较少的产出增量。这表明思政类课程大部分学习者学习效率较低,并且继续增加平台学习行为的投入也不能提升成绩。

选取不同组别效率排名分布在前、中、后的代表

表 6 思政类课程部分学习者效率及排名

决策单元	组别	综合技术效率	纯技术效率	规模效率	规模报酬趋势	超效率	排名
1022	1	1	1	1	Constant	1.86	1
653	1	1	1	1	Constant	1.00	24
712	1	0.10	0.99	0.97	Constant	0.10	521
1565	2	0.06	0.34	0.17	Decreasing	0.11	1 602

综合效率能够综合体现学习者对于学习方法的使用、学习技巧和时间管理、学习努力投入程度,该指标由纯技术效率和规模效率两部分组成,其中纯技术效率主要刻画了学习方法和时间管理等方面,规模效率则刻画了学习者在学习过程中的努力程度。进一步分析综合效率非有效的学习者的纯技术效率发现,有 889 人(占比 56.16%)纯技术效率实现有效,这部分的学习者的综合效率无效主要源于规模效率不足,需要提升努力程度。而全部学习效率非有效的学习者中规模效率实现有效的有 26 人(占比 1.64%),这部分学习者的综合效率无效主要是源于纯技术效率不足,需要在学习方法、学习技巧和时间管理等方面加强。综合效率非有效的学习者中,规模报酬递减的占比为 88.31%,这部分学习者学习投入程度不足,通过提高在线学习时间所带来的成绩提升不明显。

(2)学习效率的组间差异分析

组间差异能够反映不同学习者群体学习方法的差异性。对比两类学习者的效率指标能够发现不同的学习问题,进而提出精准的施教建议。单因素方差分析结果显示,两组学习者的效率指标在统计学上均存在显著差异( $p<0.01$ )。

综合效率值指标上,勤奋收获者均值(0.18)略小于学习体验者均值(0.27),并且在统计学上拥有较大的差异( $p<0.01$ )。规模效率值上也呈现出相同的

性决策单元进行分析,结果如表 6 所示。学习者 1022 的效率指标均实现了有效,超效率值为 1.86,排序为 1,这意味着该学习者即使再增加 86%的学习行为投入仍然可以实现有效。直观而言,学习者超效率变化趋势与综合效率变化趋势不完全相同,超效率均值为 0.87,标准差为 0.22,说明大部分学习者的超效率较好,但是距离实现有效还有一定的提升空间。超效率实现有效的学习者,超效率均值 1.10,标准差 0.21,整体超效率的波动也比较小。

数学特征,勤奋收获者的均值(0.19)略低于学习体验者(0.37),这主要源于大部分的勤奋收获者存在较大程度的投入冗余。超效率值则真实反映了两组学习者在学习效果上的差距,勤奋收获者的均值(0.93)远大于学习体验者(0.28),勤奋收获者的方差(0.004)也明显小于学习体验者(0.05),勤奋收获者的超效率组间差异更小,数据的波动更加稳定,两组超效率的方差均较小,数据之间的波动并不明显,纯技术效率反映了和超效率类似的特征。

2. 外语课程 DEA 分析结果

(1)与思政课程的效率对比分析

外语课程中实现综合效率有效的共计 14 人(占比 1.68%),同样只有少数学生实现了学习效率的有效。纯技术效率实现有效的学习者为 191 人(占比 22.87%),规模效率实现有效的为 21 人(占比 2.51%),大部分学习者的学习效率无效由规模效率不足导致,应加强努力程度。与思政课程相同,外语课程大部分学习者(占比 91.50%)也呈规模递减特征,说明学习行为投入冗余是影响学习效率的重要因素。外语课程中规模报酬不变的占比为 7.78%,也存在规模报酬递增的学习者,应加大对学习行为的投入,挖掘潜在学习能力。

整体而言,外语课程与思政课程呈现相似的整体特征:实现有效的学习者占据少数;学习效率分布不

均,外语课程综合效率小于 0.2 的学习者占比 68.86%;综合效率的均值、规模效率均值、超效率均值与纯技术效率均值呈递增分布(分别为:0.20、0.22、0.82、0.95);规模报酬递减的学习者占据多数。两类学习者在四项效率指标中均存在明显差异( $p < 0.01$ )且勤奋收获者的纯技术效率均值(0.97)明显高于学习体验者(0.82),且勤奋收获者的实现纯技术有效的比例(0.24)也明显高于学习体验者(0.14),说明在学习上表现得更优秀的学生会更重视学习技巧、时间管理等方面。

### (2)学习非有效的学习者的目标改进值

为对非有效学习者的效率改进进行定量分析,可根据 DEA 的投影定理,得到各投入产出指标的松弛改进值。产出改进对应实现有效时应提高多少产出,投入改进则表示在保持原有的效果时投入指标应减少的额外投入。产出角度的 DEA 模型,优先对产出指标进行调节再对投入指标进行调节,投入角度的 DEA 模型则相反。选取超效率值排名中间水平的决策单元 27 分析其根据 DEA 投影定理所得的松弛改进值,如表 7 所示。

表 7 外语课程决策单元 27 松弛改进情况

指标	原始值	松弛改进值	目标值
课程行为天数	10	0	10
行为总数	994	-337	657
浏览数	518	-247	271
浏览资源数	36	-33	3
试卷成绩	84	0	84
形考成绩	76	0	76
最终成绩	80	2	82

该决策单元的超效率排名为 554,超效率值 0.84,综合效率值 0.10,纯技术效率值 0.91,规模效率值 0.11 且规模报酬递减,在聚类分析时属于勤奋收获者,符合大多数决策单元的效率分布情况。由于在模型选择时选择了产出角度的模型进行了计算,因此会优先对产出指标进行优化。根据各投入产出变量所对应的松弛变量也可以发现,该学习者存在产出不足、投入冗余的情况。在行为总数、浏览数、浏览资源数三个指标上冗余值较大需要大幅度的调整,这说明课程行为总数、浏览数、浏览资源数对于学习效率的贡献值并没有预期的多,该学习者可能过度关注学

习资源的浏览但是没能对学习内容进行高效率的转化,从而导致投入过高却没有创造产出价值。存在这种情况的学习者共有 330 名,占比 39.37%,这意味着一定比例的学习者对于知识的吸收效率较低,应该是教学过程的重点关注对象。

### 3. 管理课程 DEA 分析结果

#### (1)超效率的组间差异性分析

通过对比分析发现,三类课程在效率的整体分布上具有相似特征,效率的整体趋势也较为相似。两类学习者的效率组间差异也呈相似特征,且管理课程的效率组间差异显著性也通过了检验( $p < 0.01$ )。

为探究两类学习者的效率差异,选取排序上更具优势的超效率指标,评估两类学习者的差异水平。两类学习者的超效率频数分布图 2 所示。超效率具有显著的组间差异( $p < 0.01$ ),勤奋收获者均值为 0.89,学习体验者均值为 0.13,在学习行为上投入较多的勤奋收获者超效率明显优于学习体验者,勤奋收获者方差(0.007)远小于学习体验者(0.03),成绩较好的学习者的效率分布更为稳定。由图可知,勤奋收获者的超效率主要分布在 0.9 以上(占比 79.89%),大部分勤奋收获者都属于高效的学习者,学习体验者的超效率主要分布在低于 0.2 的范围内(占比 90.38%)。勤奋收获者的高效率主要得益于纯技术效率方面,高效学习者对于学习方法的使用和时间管理能力上具有明显优势。

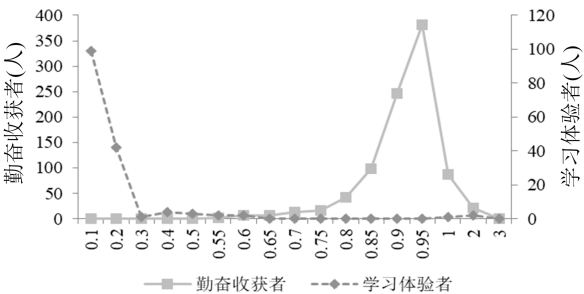


图 2 管理类课程超效率频数分布

### 四、研究结论与建议

#### (一)研究结论

当前,在线教育平台规模迅速扩展,线上学习者人数逐年递增。在此背景下,本文对在线教育平台思政、外语、管理课程的学习者学习行为数据进行挖掘,同时对学习有效性和群体差异性进行分析研究,主要研究结论如下。

第一,勤奋收获者与学习体验者这两类学习者具



有显著差异性。勤奋收获者在学习行为上的投入会明显优于学习体验者,特别是在行为总数的指标上差异最为明显。这与 De Barba 等的研究<sup>[25]</sup>发现类似,勤奋收获者在学习过程中更愿意钻研、投入更多的学习时间;而学习体验者则在学习过程中缺乏积极性,对在线课程学习往往缺少长期的学习计划。

第二,DEA 方法在不同类型的课程学习效率分析中具有普适性。三类课程的成绩及效率分布具有相似性特征,学习行为上呈现出课程行为指标组内差异明显、课程通过率高、优秀者成绩差异较小、成绩存在两极分化等现象。学习效率上,不同类型的课程均存在综合效率有效者占比较低、整体规模报酬递减的学习者占比高、纯技术效率相对规模效率较易实现有效等特征。

第三,大部分的在线学习者并未实现效率最优。就综合效率指标而言,三类课程均只存在少数实现有效的学习者,并且大部分尚未实现有效的学习者呈现出规模报酬递减特征,较多学习投入的增加也仅能实现较少的成绩提升效果。

第四,两类学习者在超效率方面差异最为显著。大部分勤奋收获者都是高效学习者,而低效学习者可能由于基础知识或适应能力不足导致学习的成果也不尽人意<sup>[26]</sup>。

## (二)策略建议

### 1. 优化教学评价体系,创新教育教学模式

传统教学评价模式下往往注重成绩的考核,然而对学习过程中的投入却难以进行量化分析,通过对于学习效率的分析可以避免教学评价指标单一的问题。进行学习效率评价,可以对学生的学习效果进行准确评判。在教学过程中,需要结合线上教学的模式,制定更加完善的筛选制度,以保证学生的学习积极性,避免资源浪费<sup>[27]</sup>。

### 2. 注重阶段性考核,提升学习效果

大量效率较低的学习者在学习过程中存在学习行次数过多、学习行为冗余的情况。在综合教学评价指标的前提下,通过阶段性考核能够及时发现高学习投入者可能存在的 learning 问题,从而督促其改进学习策略,进而提高学习效率。部分学习者也存在学习动机不足、学习行为过少等问题,导致最终没有通过考试,这部分学习者可能学习效率相对较高,但缺乏持续学习的动力支撑。教师和学习管理人员应进行阶段性监督,通过多样化的考核手段和激励方案提高学习者的学习热情。

### 3. 加强教学环节的实时干预,提升学习积极性

线上教育平台的学习自由性较高,学习者可以随时回顾教学内容,这也使得学生存在因重复性的学习活动而产生学习效率不高的问题。教师应促使学习者提升单次学习的努力程度,加强学习中时间管理的能力。在教学过程中,教师应及时提醒干预,鼓励学生加大学习的投入度从而提升学习效率。

## 参考文献:

- [1] 李学书,孙传远.在线教育治理:从野蛮生长到规范发展[J].河北师范大学学报(教育科学版),2021,23(5):80-87.
- [2] 孙发勤,冯锐.学习分析视域下的在线学习动机评估研究[J].现代教育技术,2022,32(1):94-103.
- [3] 网经社电子商务研究中心.2022年度中国数字教育市场数据报告[EB/OL].(2023-03-27)[2023-10-1].<https://www.100ec.cn/detail-6625670.html>.
- [4] 袁东斌,徐智华.MOOC在开放大学非学历继续教育中的应用研究[J].福建广播电视大学学报,2018(6):18-22.
- [5] 唐燕儿,关淑文.基于霍姆伯格远程教育思想的在线教学创新策略研究:以疫情期间成人高等教育在线教学为例[J].中国电化教育,2020(5):27-33.
- [6] 郭建鹏,张娟.我国高校智能教育的进展、问题与实现路径[J].河北师范大学学报(教育科学版),2021,23(1):63-69.
- [7] 魏顺平.Moodle平台数据挖掘研究:以一门在线培训课程学习过程分析为例[J].中国远程教育,2011(1):24-30.
- [8] 王佳利,李斌峰.基于网络教学平台校本混合课程教学效果的实证研究[J].电化教育研究,2016,37(3):101-107.
- [9] 胡祖辉,施佳.高校学生上网行为分析与数据挖掘研究[J].中国远程教育,2017(2):26-32.
- [10] 贺超凯,吴蒙.edX平台教育大数据的学习行为分析与预测[J].中国远程教育,2016(6):54-59.
- [11] WANG J, ZHANG Y. Clustering Study of Student Groups Based on Analysis of Online Learning Behavior[C]//Proceed-



- ings of the 2019 International Conference on Modern Educational Technology - ICMET 2019. Nanjing: ACM Press, 2019: 115-119.
- [12] 赵磊,邓彤,吴卓平. 基于数据挖掘的 MOOC 学习者学业成绩预测与群体特征分析[J]. 重庆高教研究, 2021,9(6): 95-105.
- [13] 王改花,傅钢善. 数据挖掘视角下网络学习者行为特征聚类分析[J]. 现代远程教育研究,2018(4): 106-112.
- [14] FUENTES R, FUSTER B, LILLO-BANULS A. A Three-stage DEA Model to Evaluate Learning-teaching Technical Efficiency: Key Performance Indicators and Contextual Variables [J]. Expert Systems with Applications, 2016(48): 89-99.
- [15] 陈长胜,刘梅,沈书生,等. 慕课学习完成者的学习效率及其群体差异性研究[J]. 中国远程教育,2021(10): 39-47.
- [16] 朱凡,王印琪. 基于 k-means 与神经网络机器学习算法的用户信息聚类及预测研究[J]. 情报科学,2021,39(7): 83-90.
- [17] 吴永生,魏巧米. 公共事业管理专业课程体系设置及学生成绩的定量分析[J]. 当代教育论坛(管理研究),2011(9): 92-94.
- [18] 成刚. 数据包络分析方法与 MaxDEA 软件[M]. 北京:知识产权出版社,2014: 24.
- [19] 鲍平平. DEA 在网络学习相对有效性评价中的应用[J]. 现代远程教育研究,2007(6): 65-68.
- [20] 廖继胜,刘昱,刘志虹,等. 基于 DEA 和截断回归模型的欠发达省份国家助学贷款绩效及其影响因素研究[J]. 黑龙江高教研究,2022,40(4): 54-60.
- [21] 王梦倩,范逸洲,郭文革,等. MOOC 学习者特征聚类分析研究综述[J]. 中国远程教育,2018(7): 9-19.
- [22] 龚艺,刘波,谭明杰,等. 远程教育中在线学习投入与学习绩效关系的实证研究[J]. 成人教育,2018,38(6): 24-28.
- [23] 沈欣忆,刘美辰,吴健伟,等. MOOC 学习者在线学习行为和学习绩效评估模型研究[J]. 中国远程教育,2020(10): 1-8.
- [24] 李梅,孙月亚,李蕾,等. 在线课程开放度对教与学行为与效果的影响:基于成人学生的研究[J]. 成人教育,2021,41(6): 18-25.
- [25] DE BARBA P G, MALEKIAN D, OLIVEIRA E A, et al. The Importance and Meaning of Session Behaviour in a MOOC [J]. Computers & Education, 2020(146): 103772.
- [26] 肖睿,刘千慧,尚俊杰,等. 学习者的学习效率评测研究:以“课工场”平台学习者的课程学习为例[J]. 现代教育技术,2021, 31(1): 62-68.
- [27] 金慧,王陈欣,罗纯源,等. 后疫情时代的高等教育:宏观趋势、关键技术与发展思考:《2021 地平线报告(教与学版)》解读 [J]. 远程教育杂志,2021,39(3): 3-10.

## Evaluation and Differentiation of Learning Efficiency of Online Education Learners

SUO Qi, ZUO Pei

(School of Economics and Management, Qingdao University of  
Science and Technology, Qingdao Shandong 266061, China)

**Abstract:** Online education has the characteristics of openness and sharing, and digging deeper into the learning behavior of online learners can help students understand the learning effect in time. Ideological and political, foreign languages and management courses on the online education platform are selected to study the learning efficiency characteristics and group difference characteristics of online learners. The research results show that the distribution of learners' behaviors and grades in different courses is similar, and most of the learners do not achieve effective learning efficiency, and there are significant differences between the learning efficiency of "diligent gainers" and "learning experiencers" who have more learning behaviors and better grades. There are obvious differences between the learning efficiency of the "diligent gainers" and the "learning experiencers". Based on this, this paper suggests that in the process of online education, we should strengthen the stage supervision, with timely intervention for students' lack of learning motivation, and improve the screening system of online education in order to improve the learners' efficiency and maximize the usage of online courses.

**Keywords:** online learning behavior; cluster analysis; data envelopment analysis; efficiency improvement strategies  
[责任编辑 许炎]